Volume 2 • Number 2 • 1999
VALUE IN HEALTH

Panel 3: Methodological Issues in Conducting Pharmacoeconomic Evaluations—Retrospective and Claims Database Studies

Co-chairs: Renee Goldberg Arnold, PharmD,[1] James G. Kotsanos, MD, MS[2]
Panelists: Brenda Motheral, RPh, PhD, MBA,[3] Scott Ramsey, MD, PhD,[4] William Crown, PhD,[5] Katherine Puder, DrPh, MS,[6] Mark Hornbrook, PhD,[7] Alan Wright, MD, MPH,[8] Michael Murray, PharmD, MPH[9]

[1]Pharmacon International, New York, NY, [2]Eli Lilly and Company, Indianapolis, IN, [3]University of Arizona, Tucson, AZ, [4]University of Washington, Seattle, WA, [5]MEDSTAT, Cambridge, MA, [6]QCORE, LLC, Philadelphia, PA, [7]Center for Health Research, Portland, OR, [8]Advanced Paradigm Inc., Hunt Valley, MD, [9]Regenstrief Institute, Indianapolis, IN

The goal of this panel was to identify key contentious methodology issues in conducting healthcare pharmacoeconomic evaluations in the context of retrospective studies and claims data. Its specific objectives were to:

- identify and prioritize the key issues associated with pharmacoeconomic and outcomes research studies using retrospective and claims data;
- identify a plan of action to resolve these issues;
- recommend next steps.

Background and Context

Pharmacoeconomic analyses can be conducted within the context of clinical studies or by using retrospective databases. Randomized clinical trials (RCTs) and observational studies, specifically retrospective database analyses, answer different questions. For example, most RCTs are designed to measure efficacy, not effectiveness. Existing databases can provide effectiveness and "real-world" data. Cost-effectiveness analyses utilizing retrospective databases can provide real-time, relevant, and comprehensive decision support.

Retrospective analyses are relatively inexpensive to perform and can be done quickly. They are reflective of specific populations that cannot be easily studied using RCTs, and for which data may be difficult to obtain. Relative to RCTs, retrospective studies tend to cover more realistic time frames and are not constrained by the limitations of a set trial period. The perspective of a particular organization's experience can be obtained, and large samples can be surveyed. Usually, the sample sizes of retrospective databases are much larger than RCTs, enhancing their statistical power to detect important differences in outcomes. Retrospective databases encompass a wealth of variables, and analyses of these data can be used for benchmarking purposes and for capturing real-world prescribing patterns. Notwithstanding the advantages of retrospective database analyses, there are challenges that face these analyses when used for health economic evaluations.

Problem Statement

Healthcare decision-makers require rapid access to information. The evidence that assists decision-makers to draw conclusions often has not been available. Both RCTs and retrospective methods using existing databases provide such information, and they typically answer different questions. Most RCTs are designed to measure efficacy, not effectiveness. "Real-world" data can be provided by database studies. The validity of retrospective analyses is often questioned because of the potential for selection bias, confounding factors, sponsorship, data quality, and privacy issues. In addition, since there may be a time lag in the availability of information about new therapies and its incorporation into and availability from databases, the timeliness of the economic evaluations resulting from a retrospective analysis can be questioned.

 1098-3015/99/$14.00/82 82–87

Issues

Eight key issues were identified, which are addressed in this document:

1. What research questions can be answered by retrospective analyses?
2. What data sources are available to answer these questions?
3. How is cost-effectiveness measured using automated databases?
4. How can data quality within a database be evaluated?
5. What types of statistical methods can be utilized to control for treatment effects?
6. What potential types of bias exist in retrospective database analyses?
7. What alternative methods for assessing selection bias are available?
8. How can transparency be ensured in retrospective database analyses?

What Research Questions Can Be Answered by Retrospective Database Analyses?

The types of economic studies that can be conducted using automated databases include cost-consequence, cost-effectiveness, and cost-of-illness analyses. Currently, cost-utility analyses are rarely addressed because of the lack of utility data captured in databases.

Of paramount importance to utilizing retrospective database analysis for health economic evaluations is the careful crafting of the research question. This question must be derived at the outset from the perspective of the appropriate parties, and such perspective may include society, the provider, the payer, or the patient.

Crafting the question is critical to the success of a project since it drives all other aspects of the research project. It dictates:

- how the literature review is performed;
- the study objectives;
- the definitions of health outcomes and variables to be studied;
- the study design, including sample size;
- the time frame to be evaluated;
- data source(s), validation, and analyses;
- the budget.

What Data Sources Are Available to Answer the Question?

The types of data sources that can be utilized for health economic evaluations based on retrospective analysis include:

- electronic medical records (integrated modules of pharmacy, laboratory, and clinical databases);
- claims data such as that from managed care databases, the Healthcare Finance Administration (HCFA), the Department of Defense, Veterans Affairs, self-insured employers, and pharmacy benefit managers (PBMs);
- encounter data such as that from a staff/group model of health maintenance organizations (HMOs);
- expert opinion;
- results of published literature, such as meta-analyses;
- patient registries;
- national survey databases, such as the Medical Expenditure Panel Survey (MEPS), formerly the National Medical Expenditure Survey (NMES).

How Is Cost-Effectiveness Measured Using Automated Databases?

Cost-effectiveness studies require comparison of two or more competing therapeutic options. It is possible to measure cost-effectiveness using automated databases. First, one must determine if the particular database is appropriate to answer the question. Evaluation depends on the disease in question and availability of the outcome measures of effectiveness. For example, it would be inappropriate to select a hospital discharge database to evaluate lithotripsy use in kidney stones if the procedure is performed on an outpatient basis. Databases containing both outcome and cost information are required to perform cost-effectiveness analyses, including, for example, administrative claims as well as laboratory and prescription data. In some cases, linkage of databases might be appropriate to address some research questions. An example of database linkage for estimating cost of care for patients with cancer has been published [1]. Outcome measures can include, for example, diabetic ketoacidosis episodes or amputations avoided, which are coded using the International Classification of Disease (ICD-9) or Current Procedural Terminology (CPT-4) codes. Intermediate outcome measures can also be used, such as the percentage reduction of low density lipoprotein cholesterol rather than occurrence of a major cardiovascular event. Currently, databases do not usually contain quality-of-life (QoL) or utility data. However, QoL data are beginning to be collected, such as in the Patient Outcomes Research Teams (PORT) databases, in selected managed care databases, and in

the managed care Health Plan Employer Data and Information Set (HEDIS) that captures data from questionnaires like the SF-36.

How Can Data Quality within a Database Be Evaluated?

Assessing the accuracy and completeness of the database [2,3] is integral to a research study based on automated database information. Data checks must be performed regularly and consistently. Points to be considered include:

- Are there missing data elements, and if so, what percentage is missing?
- Is the cohort continuously tracked over the period of interest?
- Is it possible to trace services and diagnoses across healthcare settings (e.g., hospitals, nursing homes, and clinics)?
- Were all relevant diagnoses and procedure codes accurately recorded for the visit or episode of care?
- Are data recorded uniformly using widely accepted recording standards?
- Was the accuracy of the diagnostic and utilization records verified with chart reviews or benchmarking?
- Were logical consistency checks performed, such as searching the database for illogical matches (e.g., those between "hysterectomies" and "males")?
- Are there unique identifiers for each family member?
- Were events recorded when they actually occurred?
- Is population-denominator data available? That is, does the database contain enrollment information that enables the identification of individuals without healthcare utilization?

What Types of Statistical Methods Can Be Utilized to Control for Treatment Effects?

With retrospective database studies in particular, it is important to control for as many confounding variables as possible. A need exists for statistical methods to control for these effects, which implies that multivariate analysis should be performed whenever possible.

What Potential Types of Bias Exist in Retrospective Database Analyses?

The potential areas for bias [4] within retrospective database analyses form a long list, including, but not limited to the following examples.

Selection Bias. Individuals are not typically randomly assigned to health plans or treatments. Estimates of the effects and costs can be biased due to a correlation between unobserved factors associated with treatment selection and outcomes, such as baseline health status. Sample selection bias is often referred to by other names. For example, vintage bias, which is due to variation in physician training and practice styles or in availability of technologies resulting in confounding measures of costs and outcomes, is a specific type of selection bias.

Bias from Censoring of Data. Bias can be introduced when the length of time that individuals are observed is correlated with their outcomes. For example, some studies may impose a minimum eligibility period that leads to the omission of subjects with short-term eligibility. If the reasons for the failure to meet the minimum eligibility criteria are correlated with utilization patterns, such as a death, then a biased conclusion will result. Similarly, if the length of the observation period varies for individuals, then bias may be introduced because of failure to observe utilization that occurred after the observation period ended (e.g., there is less opportunity to observe service utilization patterns for patients treated with relatively new drugs).

Measurement Error Bias. There are numerous sources of error in measurement of data that can introduce bias. For example, specialists may tend to code diagnoses with more specificity than general practitioners. Similarly, fee-for-service providers may have an incentive to document diagnoses and services relative to at-risk providers who may have an incentive to minimize the burden of documentation. Recall bias is yet another form of measurement error. Recall bias refers to the tendency, on the part of respondents, to recall service utilization that occurred more recently with greater accuracy than service utilization that occurred in the more distant past.

Misspecification Bias. As with measurement error, there are a variety of forms of misspecification errors that can lead to biased estimates. These include omitted variables, incorrect function form, and using single equation models when a multiequation model is more appropriate.

Investigator Bias/Obsolescence Bias. There are sources of bias that may be introduced that are not directly a function of retrospective data themselves. These include investigator bias, which

arises when a researcher interprets findings in the context of preconceived viewpoints or adopts a study design (e.g., exclusion criteria) that biases the study results in the direction of the researcher's preconceived viewpoints. Investigator bias is often unintentional. Similarly, obsolescence bias may occur because medical technology used during the periods covered by a retrospective database is obsolete by the time the study is conducted.

Finally, bias is not the only statistical problem that may undermine the validity of inferences drawn from retrospective studies. Other problems include the correlation of error terms among respondents (autocorrelation), nonconstant variance of error terms among respondents (heteroscedasticity), and high correlation among explanatory variables (multicollinearity).

What Alternative Methods for Assessing Selection Bias Are Available?

Perhaps the most fundamental area in which bias occurs in a retrospective database analysis is during selection of study subjects. Whereas RCTs reduce sample selection bias through randomization by evenly distributing subjects among treatment arms, retrospective database analyses are nonrandomized. Nonrandomized studies that attempt to evaluate treatment outcomes have been widely criticized [5,6] because unobserved variables might correlate with both treatment selection and outcomes. Such a correlation can result in erroneous inferences about the magnitude and statistical significance of treatment effects.

Alternative methods for assessing selection bias include using propensity scores, instrumental variables (IVs), and sample selection models; however, these methods may fail to fully control for selection bias.

Propensity score analysis has received growing attention as a methodology for reducing the bias due to inherent differences between treatment groups that go unobserved [7–9]. Although the propensity score approach is nonparametric, Angrist [10] has recently shown that propensity score analysis may be more closely related to sample selection models than previously believed.

The use of IVs in recent papers by McClellan [11,12] has been proposed to control for the confounding effects of unobserved variables. Instrumental variables are widely used by researchers to correct for a variety of statistical problems, most notably, simultaneous equation bias and errors in measurement [13].

Sample selection models attempt to control the bias introduced by unobserved variables in treatment selection, which are also correlated with the outcome variable of interest. Sample selection models have seen wide use in the econometrics literature to study labor supply decisions and to model the effectiveness of job training programs, housing programs, welfare experiments, and many others [14]. Very recently, these models have begun to find application in the health economics literature [15,16]. Selection models are a special case of IVs and may require parameter estimates.

How Can Transparency Be Ensured in Retrospective Database Analyses?

A great deal of consensus exists among the guidelines about transparency of assumptions and methods [17]. Full disclosure and detailed methodology when reporting study results of health economic evaluations is recommended by the Uniform Requirements for Manuscripts Submitted to Biomedical Journals [18], the Task Force on Principles for Economic Analysis of Healthcare Technology [19], and the Pharmaceutical Research and Manufacturers of America's Methodological and Conduct Principles for Pharmacoeconomic Research [20].

One final point is that data privacy has become an increased focus of attention. The International Society of Pharmacoeconomics and Outcomes Research (ISPOR) has recently adopted the data privacy recommendations developed by the International Society for Pharmacoepidemiology (ISPE Data Privacy).

Recommendations and Next Steps

The credibility of retrospective database analyses in health economic evaluations must be enhanced by good research design, multiple and varied checks on data quality, and attention to areas of potential bias in a given study.

The following recommendations address the issues identified:

- It is recommended that retrospective database analysis studies begin with a clear question and design, based on the International Society for Pharmacoepidemiology guidelines for good epidemiological practices [21].
- Privacy of individuals must be ensured at all times in retrospective database analyses.
- Techniques that exist to address shortcomings of retrospective data sets should be used.

- Multivariate models should be subjected to extensive specification testing.
- Examining age- or gender-adjusted utilization rates and annual per capita expense by payer, health plan, geographic region, and country is suggested.
- Administrative databases, frequently used for retrospective pharmacoeconomic studies, need to be augmented to include more clinical information (e.g., the results of lab tests, not just an indicator that a lab test was conducted).
- Bias is a threat to the validity of inferences drawn from any retrospective and claims database analysis. When bias exists, either real or perceived, standard measures should be established to deal with all areas of potential bias. Methods of addressing specific types of bias include: (1) selection bias: modeling (e.g., Heckman approach, fixed-effect, and random-effect models), propensity scores; (2) measurement error and bias: modeling (e.g., bootstrap estimation), data imputation; (3) misspecification bias and other violations of model assumptions: conduct and report specification error tests, select appropriate model and error distribution, or adjust observations to remove the specification error(s); (4) investigator bias: disclosure of conflict of interest, analysis assumptions, and study sponsor for investigator(s) and key staff; (5) obsolescence bias: use expert opinion to identify key innovative technologies and model them, look at time trends in key utilization parameters, and estimate models for separate time periods; (6) vintage bias with regard to human and physical capital: modeling, link with other data sources such as an HCFA area resource file.

Summary

Retrospective databases can be useful tools for health economic evaluations. They offer large populations using real-world information within rapid and realistic treatment periods, and can answer questions related to cost-effectiveness. Nevertheless, pharmacoeconomic studies based upon retrospective databases face a variety of threats to the validity of the inferences drawn from them. The credibility of retrospective database analyses in health economic evaluations must be enhanced by careful study design, elimination of bias, and reporting the results of these studies in a clear, believable, and transparent fashion.

References

1 Potosky A, Riley G, Lubitz J, Mentnech R, Kessler L. Potential for cancer related health services using a linked Medicare tumor registry database. Med Care 1993;31:732–48.
2 Iezzoni LI. Assessing quality using administrative data. Ann Intern Med 1997;127(8 Part 2):665–74.
3 The MEDSTAT Group. A guide for states to assist in the collection and analysis of Medicaid managed care data. (Prepared for the Health Care Financing Administration.) No. 500–92–0035. 1997.
4 Sackett D. Bias in analytic research. J Chronic Dis 1979;32:51–63.
5 Andersson F. Why is the pharmaceutical industry investing increasing amounts in health economic evaluations? Int J Tech Ass Health Care 1995; 11:750–61.
6 Canadian Coordinating Office for Health Technology Assessment. Guidelines for Economic Evaluation of Pharmaceuticals. Ottawa: Canadian Coordinating Office for Health Technology Assessment, 1997.
7 Drake C, Fisher L. Prognostic models and the propensity score. Int J Epidemiol 1995;24:183–7.
8 Robins J, Mark S, Newey W. Estimating exposure effects by modelling the expectation of exposure conditional on confounders. Biometrics 1992; 48: 479–95.
9 Rosenbaum P, Rubin D. Reducing bias in observational studies using subclassification on the propensity score. J Am Statist Assoc 1984;79:516–24.
10 Angrist J. Conditional independence in sample selection models. Economic Lett 1997;54:103–12.
11 McClellan M, McNeil B, Newhouse J. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA 1994;272:859–66.
12 McClellan M. Uncertainty, health-care technologies, and health-care choices. The Economics of Health and Health Care 1995;85(2):38–44.
13 Kennedy P. Violating assumption four: measurement errors and autoregression. In: A Guide to Econometrics (3rd ed.). Cambridge, MA: The MIT Press 1992:134–50.
14 Heckman J, Smith J. Assessing the case for social experiments. J Economic Perspectives 1995;9:85–110.
15 Dowd B, Feldman R, Moscovice I, Wisner C, Bland P, Finch M. An analysis of selectivity bias in the Medicare AAPCC. Health Care Financing Rev 1996;17:35–57.
16 Hylan T, Kotsanos J, Anderson J, Brown S, Copley-Merriman C, Egbuonu-Davis L, et al. Comparison of a decision analytic model with results from a naturalistic economic clinical trial: an application to evaluating alternative antidepressants. Am J Managed Care 1996;2:1211–23.

17 Mullins CD. An Update of Pharmacoeconomic Guidelines and Principles. Presented at the ISPOR Conference, Issues in Pharmacoeconomics, Crystal City, VA, 18 February, 1998.
18 International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. New Eng J Med 1997;336:309–15.
19 Task Force on Principles for Economic Analysis of Health Care Technology. Economic analysis of healthcare technology: a report on principles. Ann Intern Med 1995;123:61–70.
20 Pharmaceutical Research and Manufacturers of America. Methodological and conduct principles for pharmacoeconomic research. 1995.
21 International Society for Pharmacoepidemiology. Good research practice guidelines. Pharmacoepidemiol Drug Safety 1996;5:333–8.